SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: HUSE, WILLIAM D.

(ii) TITLE OF INVENTION: SURFACE EXPRESSION LIBRARIES OF HETEROMERIC RECEPTORS

(iii) NUMBER OF SEQUENCES: 75

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
(B) STREET: 444 SO. FLOWER STREET, SUITE 200
(C) CITY: LOS ANGELES
(D) STATE: CALIFORNIA
(E) COUNTRY: UNITED STATES
(F) ZIP: 90071

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: CAMPBELL, CATHRYN A.
(B) REGISTRATION NUMBER: 31,815
(C) REFERENCE/DOCKET NUMBER: P31 8882

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 619-535-9001
(B) TELEFAX: 619-535-8949

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7445 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAT 60

ATAGCTAAAC AGGTTATTGA CCATTTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT 120

CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA 180

GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA 240

TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG 300

TTGGAGTTTG CTTCCGGTCT GGTTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG 360

TCTTTCGGGC TTCCTCTTAA TCTTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT 420

```
CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA   480
TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT   540
AAACATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT   600
GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT   660
AATTCCTTTT GGCGTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG   720
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT   780
TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA   840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT   900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG   960
AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC  1020
TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC  1080
GTCTGCGCCT CGTTCCGGCT AAGTAACATG GAGCAGGTCG CGGATTTCGA CACAATTTAT  1140
CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT  1200
CAAAGATGAG TGTTTTAGTG TATTCTTTCG CCTCTTTCGT TTTAGGTTGG TGCCTTCGTA  1260
GTGGCATTAC GTATTTTACC CGTTTAATGG AAACTTCCTC ATGAAAAAGT CTTTAGTCCT  1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA  1380
CGATCCCGCA AAAGCGGCCT TTAACTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA  1440
TGCGTGGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTTAAGAA  1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT  1560
TTTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTCGCAA TTCCTTTAGT TGTTCCTTTC  1620
TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA  1680
TTTACTAACG TCTGGAAAGA CGACAAAACT TTAGATCGTT ACGCTAACTA TGAGGGTTGT  1740
CTGTGGAATG CTACAGGCGT TGTAGTTTGT ACTGGTGACG AAACTCAGTG TTACGGTACA  1800
TGGGTTCCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT  1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT  1920
ATTCCGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA  1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTTT  2040
CAGAATAATA GGTTCCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT  2100
CAAGGCACTG ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG  2160
TATGACGCTT ACTGGAACGG TAAATTCAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA  2220
GATCCATTCG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT  2280
GCTGGCGGCG GCTCTGGTGG TGGTTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT  2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTC TGGTTCCGGT  2400
GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT  2460
```

GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACTTGATT CTGTCGCTAC TGATTACGGT 2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT 2580
GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT 2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT 2700
TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA 2760
TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG 2820
TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT 2880
TATTATTGCG TTTCCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC 2940
TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CTATTTCATT GTTTGTTGCT CTTATTATTG 3000
GGCTTAACTC AATTCTTGTG GGTTATCTCT CTGATATTAG CGCTCAATTA CCCTCTGACT 3060
TTGTTCAGGG TGTTCAGTTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATTC 3120
TCTCTGTAAA GGCTGCTATT TTCATTTTTG ACGTTAAACA AAAAATCGTT TCTTATTTGG 3180
ATTGGGATAA ATAATATGGC TGTTTATTTT GTAACTGGCA AATTAGGCTC TGGAAAGACG 3240
CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT 3300
CTTGATTTAA GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAAC GCCTCGCGTT 3360
CTTAGAATAC CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGGCG CGGTAATGAT 3420
TCCTACGATG AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTGCGGTAC TTGGTTTAAT 3480
ACCCGTTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTTCT ACATGCTCGT 3540
AAATTAGGAT GGGATATTAT TTTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG 3600
CGTTCTGCAT TAGCTGAACA TGTTGTTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT 3660
TTTGTCGGTA CTTTATATTC TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT 3720
GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT 3780
ACTGGTAAGA ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTTCTAG TAATTATGAT 3840
TCCGGTGTTT ATTCTTATTT AACGCCTTAT TTATCACACG GTCGGTATTT CAAACCATTA 3900
AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTTGA AAAAGTTTTC ACGCGTTCTT 3960
TGTCTTGCGA TTGGATTTGC ATCAGCATTT ACATATAGTT ATATAACCCA ACCTAAGCCG 4020
GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTTGATA AATTCACTAT TGACTCTTCT 4080
CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT 4140
AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTTATG TACTGTTTCC 4200
ATTAAAAAAG GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTTGTTT TCTTGATGTT 4260
TGTTTCATCA TCTTCTTTTG CTCAGGTAAT TGAAATGAAT AATTCGCCTC TGCGCGATTT 4320
TGTAACTTGG TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTTCTCCCG ATGTAAAAGG 4380
TACTGTTACT GTATATTCAT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTTATTTC 4440
TGTTTTACGT GCTAATAATT TTGATATGGT TGGTTCAATT CCTTCCATAA TTCAGAAGTA 4500

TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA 4560
TGATAATTCC GCTCCTTCTG GTGGTTTCTT TGTTCCGCAA AATGATAATG TTACTCAAAC 4620
TTTTAAAATT AATAACGTTC GGGCAAAGGA TTTAATACGA GTTGTCGAAT TGTTTGTAAA 4680
GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT 4740
TAGTGCACCT AAAGATATTT TAGATAACCT TCCTCAATTC CTTTCTACTG TTGATTTGCC 4800
AACTGACCAG ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA 4860
TTTTTCATTT GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG 4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTCGGT ATTTTTAATG GCGATGTTTT 4980
AGGGCTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG 5040
TATTCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTTAT 5100
TACTGGTCGT GTGACTGGTG AATCTGCCAA TGTAAATAAT CCATTTCAGA CGATTGAGCG 5160
TCAAAATGTA GGTATTTCCA TGAGCGTTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT 5220
TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT 5280
TACTAATCAA AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTTACT 5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA 5400
AATCCCTTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT 5460
ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG 5520
GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT 5580
TCGCTTTCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC 5640
GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG 5700
ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA 5760
CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTTCCA AACTGGAACA ACACTCAACC 5820
CTATCTCGGG CTATTCTTTT GATTTATAAG GGATTTTGCC GATTTCGGAA CCACCATCAA 5880
ACAGGATTTT CGCCTGCTGG GGCAAACCAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG 5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCCT 6000
GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC 6060
ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC 6120
TCACTCATTA GGCACCCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA 6180
TTGTGAGCGG ATAACAATTT CACACGCGTC ACTTGGCACT GGCCGTCGTT TTACAACGTC 6240
GTGACTGGGA AAACCCTGGC GTTACCCAAG CTTTGTACAT GGAGAAAATA AAGTGAAACA 6300
AAGCACTATT GCACTGGCAC TCTTACCGTT ACCGTTACTG TTTACCCCTG TGACAAAAGC 6360
CGCCCAGGTC CAGCTGCTCG AGTCAGGCCT ATTGTGCCCA GGGGATTGTA CTAGTGGATC 6420
CTAGGCTGAA GGCGATGACC CTGCTAAGGC TGCATTCAAT AGTTTACAGG CAAGTGCTAC 6480
TGAGTACATT GGCTACGCTT GGGCTATGGT AGTAGTTATA GTTGGTGCTA CCATAGGGAT 6540

TAAATTATTC AAAAAGTTTA CGAGCAAGGC TTCTTAAGCA ATAGCGAAGA GGCCCGCACC 6600

GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT GGCGCTTTGC CTGGTTTCCG 6660

GCACCAGAAG CGGTGCCGGA AAGCTGGCTG GAGTGCGATC TTCCTGAGGC CGATACGGTC 6720

GTCGTCCCCT CAAACTGGCA GATGCACGGT TACGATGCGC CCATCTACAC CAACGTAACC 6780

TATCCCATTA CGGTCAATCC GCCGTTTGTT CCCACGGAGA ATCCGACGGG TTGTTACTCG 6840

CTCACATTTA ATGTTGATGA AAGCTGGCTA CAGGAAGGCC AGACGCGAAT TATTTTTGAT 6900

GGCGTTCCTA TTGGTTAAAA AATGAGCTGA TTTAACAAAA ATTTAACGCG AATTTTAACA 6960

AAATATTAAC GTTTACAATT TAAATATTTG CTTATACAAT CTTCCTGTTT TTGGGGCTTT 7020

TCTGATTATC AACCGGGGTA CATATGATTG ACATGCTAGT TTTACGATTA CCGTTCATCG 7080

ATTCTCTTGT TTGCTCCAGA CTCTCAGGCA ATGACCTGAT AGCCTTTGTA GATCTCTCAA 7140

AAATAGCTAC CCTCTCCGGC ATTAATTTAT CAGCTAGAAC GGTTGAATAT CATATTGATG 7200

GTGATTTGAC TGTCTCCGGC CTTTCTCACC CTTTTGAATC TTTACCTACA CATTACTCAG 7260

GCATTGCATT TAAAATATAT GAGGGTTCTA AAAATTTTTA TCCTTGCGTT GAAATAAAGG 7320

CTTCTCCCGC AAAAGTATTA CAGGGTCATA ATGTTTTTGG TACAACCGAT TTAGCTTTAT 7380

GCTCTGAGGC TTTATTGCTT AATTTTGCTA ATTGTTTGCC TTGCCTGTAT GATTTATTGG 7440

ACGTT 7445

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7317 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAT 60

ATAGCTAAAC AGGTTATTGA CCATTTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT 120

CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA 180

GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA 240

TCCGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG 300

TTGGAGTTTG CTTCCGGTCT GGTTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG 360

TCTTTCGGGC TTCCTCTTAA TCTTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT 420

CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA 480

TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT 540

AAACATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT 600

GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT 660

AATTCCTTTT GGCGTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG 720

```
ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT   780
TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA   840
CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT   900
CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG   960
AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC  1020
TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGAGC  1080
GTCTGCGCCT CGTTCCGGCT AAGTAACATG GAGCAGGTCG CGGATTTCGA CACAATTTAT  1140
CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT  1200
CAAAGATGAG TGTTTTAGTG TATTCTTTCG CCTCTTTCGT TTTAGGTTGG TGCCTTCGTA  1260
GTGGCATTAC GTATTTTACC CGTTTAATGG AAACTTCCTC ATGAAAAAGT CTTTAGTCCT  1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA  1380
CGATCCCGCA AAAGCGGCCT TTAACTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA  1440
TGCGTGGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTTAAGAA  1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT  1560
TTTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTCGCAA TTCCTTTAGT TGTTCCTTTC  1620
TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA  1680
TTTACTAACG TCTGGAAAGA CGACAAAACT TTAGATCGTT ACGCTAACTA TGAGGGTTGT  1740
CTGTGGAATG CTACAGGCGT TGTAGTTTGT ACTGGTGACG AAACTCAGTG TTACGGTACA  1800
TGGGTTCCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT  1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT  1920
ATTCCGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA  1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTTT  2040
CAGAATAATA GGTTCCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT  2100
CAAGGCACTG ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG  2160
TATGACGCTT ACTGGAACGG TAAATTCAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA  2220
GATCCATTCG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT  2280
GCTGGCGGCG GCTCTGGTGG TGGTTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT  2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTC TGGTTCCGGT  2400
GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT  2460
GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACTTGATT CTGTCGCTAC TGATTACGGT  2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT  2580
GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT  2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT  2700
TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA  2760
```

TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG 2820

TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT 2880

TATTATTGCG TTTCCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC 2940

TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CTATTTCATT GTTTCTTGCT CTTATTATTG 3000

GGCTTAACTC AATTCTTGTG GGTTATCTCT CTGATATTAG CGCTCAATTA CCCTCTGACT 3060

TTGTTCAGGG TGTTCAGTTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATTC 3120

TCTCTGTAAA GGCTGCTATT TTCATTTTTG ACGTTAAACA AAAAATCGTT TCTTATTTGG 3180

ATTGGGATAA ATAATATGGC TGTTTATTTT GTAACTGGCA AATTAGGCTC TGGAAAGACG 3240

CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT 3300

CTTGATTTAA GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAAC GCCTCGCGTT 3360

CTTAGAATAC CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGGCG CGGTAATGAT 3420

TCCTACGATG AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTGCGGTAC TTGGTTTAAT 3480

ACCCGTTCTT GGAATGATAA GGAAAGACAG GCGATTATTG ATTGGTTTCT ACATGCTCGT 3540

AAATTAGGAT GGGATATTAT TTTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG 3600

CGTTCTGCAT TAGCTGAACA TGTTGTTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT 3660

TTTGTCGGTA CTTTATATTC TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT 3720

GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT 3780

ACTGGTAAGA ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTTCTAG TAATTATGAT 3840

TCCGGTGTTT ATTCTTATTT AACGCCTTAT TTATCACACG GTCGGTATTT CAAACCATTA 3900

AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTTGA AAAAGTTTTC ACGCGTTCTT 3960

TGTCTTGCGA TTGGATTTGC ATCAGCATTT ACATATAGTT ATATAACCCA ACCTAAGCCG 4020

GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTTGATA AATTCACTAT TGACTCTTCT 4080

CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT 4140

AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTTATG TACTGTTTCC 4200

ATTAAAAAAG GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTTGTTT TCTTGATGTT 4260

TGTTTCATCA TCTTCTTTTG CTCAGGTAAT TGAAATGAAT AATTCGCCTC TGCGCGATTT 4320

TGTAACTTGG TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTTCTCCCG ATGTAAAAGG 4380

TACTGTTACT GTATATTCAT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTTATTTC 4440

TGTTTTACGT GCTAATAATT TTGATATGGT TGGTTCAATT CCTTCCATAA TTCAGAAGTA 4500

TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA 4560

TGATAATTCC GCTCCTTCTG GTGGTTTCTT TGTTCCGCAA AATGATAATG TTACTCAAAC 4620

TTTTAAAATT AATAACGTTC GGGCAAAGGA TTTAATACGA GTTGTCGAAT TGTTTGTAAA 4680

GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT 4740

TAGTGCACCT AAAGATATTT TAGATAACCT TCCTCAATTC CTTTCTACTG TTGATTTGCC 4800

AACTGACCAG ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA 4860
TTTTTCATTT GCTGCTGGGT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG 4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTCGGT ATTTTTAATG GCGATGTTTT 4980
AGGGCTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG 5040
TATTCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTTAT 5100
TACTGGTCGT GTGACTGGTG AATCTGCCAA TGTAAATAAT CCATTTCAGA CGATTGAGCG 5160
TCAAAATGTA GGTATTTCCA TGAGCGTTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT 5220
TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT 5280
TACTAATCAA AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTTACT 5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA 5400
AATCCCTTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT 5460
ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG 5520
GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT 5580
TCGCTTTCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC 5640
GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG 5700
ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA 5760
CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTTCCA AACTGGAACA ACACTCAACC 5820
CTATCTCGGG CTATTCTTTT GATTTATAAG GGATTTTGCC GATTTCGGAA CCACCATCAA 5880
ACAGGATTTT CGCCTGCTGG GGCAAACCAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG 5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCCT 6000
GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC 6060
ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC 6120
TCACTCATTA GGCACCCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA 6180
TTGTGAGCGG ATAACAATTT CACACGCCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC 6240
TACGGCAGCC GCTGGATTGT TATTACTCGC TGCCCAACCA GCCATGGCCG AGCTCGTGAT 6300
GACCCAGACT CCAGATATCC AACAGGAATG AGTGTTAATT CTAGAACGCG TCACTTGGCA 6360
CTGGCCGTCG TTTTACAACG TCGTGACTGG GAAAACCCTG GCGTTACCCA AGCTTAATCG 6420
CCTTGCAGAA TTCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC 6480
TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCTTT GCCTGGTTTC CGGCACCAGA 6540
AGCGGTGCCG GAAAGCTGGC TGGAGTGCGA TCTTCCTGAG GCCGATACGG TCGTCGTCCC 6600
CTCAAACTGG CAGATGCACG GTTACGATGC GCCCATCTAC ACCAACGTAA CCTATCCCAT 6660
TACGGTCAAT CCGCCGTTTG TTCCCACGGA GAATCCGACG GGTTGTTACT CGCTCACATT 6720
TAATGTTGAT GAAAGCTGGC TACAGGAAGG CCAGACGCGA ATTATTTTTG ATGGCGTTCC 6780
TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTTTAA CAAAATATTA 6840

ACGTTTACAA TTTAAATATT TGCTTATACA ATCTTCCTGT TTTTGGGGCT TTTCTGATTA 6900

TCAACCGGGG TACATATGAT TGACATGCTA GTTTTACGAT TACCGTTCAT CGATTCTCTT 6960

GTTTGCTCCA GACTCTCAGG CAATGACCTG ATAGCCTTTG TAGATCTCTC AAAAATAGCT 7020

ACCCTCTCCG GCATTAATTT ATCAGCTAGA ACGGTTGAAT ATCATATTGA TGGTGATTTG 7080

ACTGTCTCCG GCCTTTCTCA CCCTTTTGAA TCTTTACCTA CACATTACTC AGGCATTGCA 7140

TTTAAAATAT ATGAGGGTTC TAAAAATTTT TATCCTTGCG TTGAAATAAA GGCTTCTCCC 7200

GCAAAAGTAT TACAGGGTCA TAATGTTTTT GGTACAACCG ATTTAGCTTT ATGCTCTGAG 7260

GCTTTATTGC TTAATTTTGC TAATTCTTTG CCTTGCCTGT ATGATTTATT GGATGTT 7317

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7729 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAT 60

ATAGCTAAAC AGGTTATTGA CCATTTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT 120

CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA 180

GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA 240

TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG 300

TTGGAGTTTG CTTCCGGTCT GGTTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG 360

TCTTTCGGGC TTCCTCTTAA TCTTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT 420

CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA 480

TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT 540

AAACATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT 600

GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT 660

AATTCCTTTT GGCGTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG 720

ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT 780

TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA 840

CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT 900

CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG 960

AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC 1020

TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC 1080

GTCTGCGCCT CGTTCCGGCT AAGTAACATG GAGCAGGTCG CGGATTTCGA CACAATTTAT 1140

CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT 1200

CAAAGATGAG TGTTTTAGTG TATTCTTTCG CCTCTTTCGT TTTAGGTTGG TGCCTTCGTA 1260
GTGGCATTAC GTATTTTACC CGTTTAATGG AAACTTCCTC ATGAAAAAGT CTTTAGTCCT 1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA 1380
CGATCCCGCA AAAGCGGCCT TTAACTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA 1440
TGCGTGGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTTAAGAA 1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT 1560
TTTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTCGCAA TTCCTTTAGT TGTTCCTTTC 1620
TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA 1680
TTTACTAACG TCTGGAAAGA CGACAAAACT TTAGATCGTT ACGGTAACTA TGAGGGTTGT 1740
CTGTGGAATG CTACAGGCGT TGTAGTTTGT ACTGGTGACG AAACTCAGTG TTACGGTACA 1800
TGGGTTCCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT 1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT 1920
ATTCCGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA 1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTTT 2040
CAGAATAATA GGTTCCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT 2100
CAAGGCACTG ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG 2160
TATGACGCTT ACTGGAACGG TAAATTCAGA GACTGCGCTT TCCATTCTGG CTTTAATGAA 2220
GATCCATTCG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT 2280
GCTGGCGGCG GCTCTGGTGG TGGTTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT 2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTC TGGTTCCGGT 2400
GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT 2460
GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACTTGATT CTGTCGCTAC TGATTACGGT 2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT 2580
GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT 2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT 2700
TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA 2760
TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG 2820
TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT 2880
TATTATTGCG TTTCCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC 2940
TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CTATTTCATT GTTTCTTGCT CTTATTATTG 3000
GGCTTAACTC AATTCTTGTG GGTTATCTCT CTGATATTAG CGCTCAATTA CCCTCTGACT 3060
TTGTTCAGGG TGTTCAGTTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATTC 3120
TCTCTGTAAA GGCTGCTATT TTCATTTTTG ACGTTAAACA AAAAATCGTT TCTTATTTGG 3180
ATTGGGATAA ATAATATGGC TGTTTATTTT GTAACTGGCA AATTAGGCTC TGGAAAGACG 3240

```
CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT   3300
CTTGATTTAA GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAAC GCCTCGCGTT   3360
CTTAGAATAC CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGGCG CGGTAATGAT   3420
TCCTACGATG AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTGCGGTAC TTGGTTTAAT   3480
ACCCGTTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTTCT ACATGCTCGT   3540
AAATTAGGAT GGGATATTAT TTTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG   3600
CGTTCTGCAT TAGCTGAACA TGTTGTTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT   3660
TTTGTCGGTA CTTTATATTC TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT   3720
GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT   3780
ACTGGTAAGA ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTTCTAG TAATTATGAT   3840
TCCGGTGTTT ATTCTTATTT AACGCCTTAT TTATCACACG GTCGGTATTT CAAACCATTA   3900
AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTTGA AAAAGTTTTC ACGCGTTCTT   3960
TGTCTTGCGA TTGGATTTGC ATCAGCATTT ACATATAGTT ATATAACCCA ACCTAAGCCG   4020
GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTTGATA AATTCACTAT TGACTCTTCT   4080
CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT   4140
AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTTATG TACTGTTTCC   4200
ATTAAAAAAG GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTTGTTT TCTTGATGTT   4260
TGTTTCATCA TCTTCTTTTG CTCAGGTAAT TGAAATGAAT AATTCGCCTC TGCGCGATTT   4320
TGTAACTTGG TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTTCTCCCG ATGTAAAAGG   4380
TACTGTTACT GTATATTCAT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTTATTTC   4440
TGTTTTACGT GCTAATAATT TTGATATGGT TGGTTCAATT CCTTCCATAA TTCAGAAGTA   4500
TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA   4560
TGATAATTCC GCTCCTTCTG GTGGTTTCTT TGTTCCGCAA AATGATAATG TTACTCAAAC   4620
TTTTAAAATT AATAACGTTC GGGCAAAGGA TTTAATACGA GTTGTCGAAT TGTTTGTAAA   4680
GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT   4740
TAGTGCACCT AAAGATATTT TAGATAACCT TCCTCAATTC CTTTCTACTG TTGATTTGCC   4800
AACTGACCAG ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA   4860
TTTTTCATTT GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG   4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTCGGT ATTTTTAATG GCGATGTTTT   4980
AGGGCTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG   5040
TATTCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTTAT   5100
TACTGGTCGT GTGACTGGTG AATCTGCCAA TGTAAATAAT CCATTTCAGA CGATTGAGCG   5160
TCAAAATGTA GGTATTTCCA TGAGCGTTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT   5220
TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT   5280
```

```
TACTAATCAA AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTTACT    5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA    5400
AATCCCTTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT    5460
ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG    5520
GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT    5580
TCGCTTTCTT CCCTTCCTTT CTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC    5640
GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG    5700
ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA    5760
CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTTCCA AACTGGAACA ACACTCAACC    5820
CTATCTCGGG CTATTCTTTT GATTTATAAG GGATTTTGCC GATTTCGGAA CCACCATCAA    5880
ACAGGATTTT CGCCTGCTGG GGCAAACCAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG    5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGGTG GTGAAAAGAA AAACCACCCT    6000
GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC    6060
ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC    6120
TCACTCATTA GGCACCCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA    6180
TTGTGAGCGG ATAACAATTT CACACGCGTC ACTTGGCACT GGCCGTCGTT TTACAACGTC    6240
GTGACTGGGA AAACCCTGGC GTTACCCAAG CTTTGTACAT GGAGAAAATA AAGTGAAACA    6300
AAGCACTATT GCACTGGCAC TCTTACCGTT ACTGTTTACC CCTGTGGCAA AAGCCCAGGT    6360
CCAGCTGCTC GAGTCGGTCT TCCCCCTGGC ACCCTCCTCC AAGAGCACCT CTGGGGGCAC    6420
AGCGGCCCTG GGCTGCCTGG TCAAGACTAA TTCCCCGAAC CGGTGACGGT GTCGTGGAAC    6480
TCAGGCGCCC TGACCAGCGG CGTGCACACC TTCCCGGCTG TCCTACAGTC CTCAGGACTC    6540
TACTCCCTCA GCAGCGTGGT GACCGTGCCC TCCAGCAGCT TGGGCACCCA GACCTACATC    6600
TGCAACGTGA ATCACAAGCC CAGCAACACC AAGGTGGACA AGAAAGCAGA GCCCAAATCT    6660
TGTACTAGTG GATCCTACCC GTACGACGTT CCGGACTACG CTTCTTAGGC TGAAGGCGAT    6720
GACCCTGCTA AGGCTGCATT CAATAGTTTA CAGGCAAGTG CTACTGAGTA CATTGGCTAC    6780
GCTTGGGCTA TGGTAGTAGT TATAGTTGGT GCTACCATAG GGATTAAATT ATTCAAAAAG    6840
TTTACGAGCA AGGCTTCTTA AGCAATAGCG AAGAGGCCCG CACCGATCGC CCTTCCCAAC    6900
AGTTGCGCAG CCTGAATGGC GAATGGCGCT TTGCCTGGTT TCCGGCACCA GAAGCGGTGC    6960
CGGAAAGCTG GCTGGAGTGC GATCTTCCTG AGGCCGATAC GGTCGTCGTC CCCTCAAACT    7020
GGCAGATGCA CGGTTACGAT GCGCCCATCT ACACCAACGT AACCTATCCC ATTACGGTCA    7080
ATCCGCCGTT TGTTCCCACG GAGAATCCGA CGGGTTGTTA CTCGCTCACA TTTAATGTTG    7140
ATGAAAGCTG GCTACAGGAA GGCCAGACGC GAATTATTTT TGATGGCGTT CCTATTGGTT    7200
AAAAAATGAG CTGATTTAAC AAAAATTTAA CGCGAATTTT AACAAAATAT TAACGTTTAC    7260
AATTTAAATA TTTGCTTATA CAATCTTCCT GTTTTTGGGG CTTTTCTGAT TATCAACCGG    7320
```

GGTACATATG ATTGACATGC TAGTTTTACG ATTACCGTTC ATCGATTCTC TTGTTTGCTC 7380

CAGACTCTCA GGCAATGACC TGATAGCCTT TGTAGATCTC TCAAAAATAG CTACCCTCTC 7440

CGGCATTAAT TTATCAGCTA GAACGGTTGA ATATCATATT GATGGTGATT TGACTGTCTC 7500

CGGCCTTTCT CACCCTTTTG AATCTTTACC TACACATTAC TCAGGCATTG CATTTAAAAT 7560

ATATGAGGGT TCTAAAAATT TTTATCCTTG CGTTGAAATA AAGGCTTCTC CCGCAAAAGT 7620

ATTACAGGGT CATAATGTTT TTGGTACAAC CGATTTAGCT TTATGCTCTG AGGCTTTATT 7680

GCTTAATTTT GCTAATTCTT TGCCTTGCCT GTATGATTTA TTGGACGTT 7729

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7557 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAT 60

ATAGCTAAAC AGGTTATTGA CCATTTGCGA AATGTATCTA ATGGTCAAAC TAAATCTACT 120

CGTTCGCAGA ATTGGGAATC AACTGTTACA TGGAATGAAA CTTCCAGACA CCGTACTTTA 180

GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA 240

TCCGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG 300

TTGGAGTTTG CTTCCGGTCT GGTTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG 360

TCTTTCGGGC TTCCTCTTAA TCTTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT 420

CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA 480

TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT 540

AAACATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT 600

GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT 660

AATTCCTTTT GGCGTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG 720

ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT 780

TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA 840

CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT 900

CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG 960

AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC 1020

TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC 1080

GTCTGCGCCT CGTTCCGGCT AAGTAACATG GAGCAGGTCG CGGATTTCGA CACAATTTAT 1140

CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT 1200

CAAAGATGAG TGTTTTAGTG TATTCTTTCG CCTCTTTCGT TTTAGGTTGG TGCCTTCGTA 1260

GTGGCATTAC GTATTTTACC CGTTTAATGG AAACTTCCTC ATGAAAAAGT CTTTAGTCCT 1320
CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA 1380
CGATCCCGCA AAAGCGGCCT TTAACTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA 1440
TGCGTGGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTTAAGAA 1500
ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT 1560
TTTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTCGCAA TTCCTTTAGT TGTTCCTTTC 1620
TATTCTCACT CCGCTGAAAC TGTTGAAAGT GTTTAGCAA AACCCCATAC AGAAAATTCA 1680
TTTACTAACG TCTGGAAAGA CGACAAAACT TTAGATCGTT ACGCTAACTA TGAGGGTTGT 1740
CTGTGGAATG CTACAGGCGT TGTAGTTTGT ACTGGTGACG AAACTCAGTG TTACGGTACA 1800
TGGGTTCCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT 1860
TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT 1920
ATTCCGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA 1980
AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTTT 2040
CAGAATAATA GGTTCCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT 2100
CAAGGCACTG ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG 2160
TATGACGCTT ACTGGAACGG TAAATTCAGA CACTGCGCTT TCCATTCTGG CTTTAATGAA 2220
GATCCATTCG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT 2280
GCTGGCGGCG GCTCTGGTGG TGGTTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT 2340
GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTC TGGTTCCGGT 2400
GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT 2460
GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACTTGATT CTGTCGCTAC TGATTACGGT 2520
GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT 2580
GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT 2640
TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT 2700
TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA 2760
TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG 2820
TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT 2880
TATTATTGCG TTTCCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC 2940
TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CGTGTTTCTT GCTCTTATTA TTGGGCTTAA 3000
CTCAATTCTT GTGGGTTATC TCTCTGATAT TAGCGCTCAA TTACCCTCTG ACTTTGTTCA 3060
GGGTGTTCAG TTAATTCTCC CGTCTAATGC GCTTCCCTGT TTTTATGTTA TTCTCTCTGT 3120
AAAGGCTGCT ATTTTCATTT TTGACGTTAA ACAAAAAATC GTTTCTTATT TGGATTGGGA 3180
TAAATAATAT GGCTGTTTAT TTTGTAACTG GCAAATTAGG CTCTGGAAAG ACGCTCGTTA 3240
GCGTTGGTAA GATTCAGGAT AAAATTGTAG CTGGGTGCAA AATAGCAACT AATCTTGATT 3300

TAAGGCTTCA AAAGCTCCCG CAAGTCGGGA GGTTCGCTAA AACGCCTCGC GTTCTTAGAA 3360
TACCGGATAA GCCTTCTATA TCTGATTTGC TTGCTATTGG GCGCGGTAAT GATTCCTACG 3420
ATGAAAATAA AAACGGCTTG CTTGTTCTCG ATGAGTGCGG TACTTGGTTT AATACCCGTT 3480
CTTGGAATGA TAAGGAAAGA CAGCCGATTA TTGATTGGTT TCTACATGCT CGTAAATTAG 3540
GATGGGATAT TATTTTTCTT GTTCAGGACT TATCTATTGT TGATAAACAG GCGCGTTCTG 3600
CATTAGCTGA ACATGTTGTT TATTGTCGTC GTCTGGACAG AATTACTTTA CCTTTTGTCG 3660
GTACTTTATA TTCTCTTATT ACTGGCTCGA AAATGCCTCT GCCTAAATTA CATGTTGGCG 3720
TTGTTAAATA TGGCGATTCT CAATTAAGCC CTACTGTTGA GCGTTGGCTT TATACTGGTA 3780
AGAATTTGTA TAACGCATAT GATACTAAAC AGGCTTTTTC TAGTAATTAT GATTCCGGTG 3840
TTTATTCTTA TTTAACGCCT TATTTATCAC ACGGTCGGTA TTTCAAACCA TTAAATTTAG 3900
GTCAGAAGAT GAAGCTTACT AAAATATATT TGAAAAAGTT TTCACGCGTT CTTTGTCTTG 3960
CGATTGGATT TGCATCAGCA TTTACATATA GTTATATAAC CCAACCTAAG CCGGAGGTTA 4020
AAAAGGTAGT CTCTCAGACC TATGATTTTG ATAAATTCAC TATTGACTCT TCTCAGCGTC 4080
TTAATCTAAG CTATCGCTAT GTTTTCAAGG ATTCTAAGGG AAAATTAATT AATAGCGACG 4140
ATTTACAGAA GCAAGGTTAT TCACTCACAT ATATTGATTT ATGTACTGTT TCCATTAAAA 4200
AAGGTAATTC AAATGAAATT GTTAAATGTA ATTAATTTTG TTTTCTTGAT GTTTGTTTCA 4260
TCATCTTCTT TTGCTCAGGT AATTGAAATG AATAATTCGC CTCTGCGCGA TTTTGTAACT 4320
TGGTATTCAA AGCAATCAGG CGAATCCGTT ATTGTTTCTC CCGATGTAAA AGGTACTGTT 4380
ACTGTATATT CATCTGACGT TAAACCTGAA AATCTACGCA ATTTCTTTAT TTCTGTTTTA 4440
CGTGCTAATA ATTTTGATAT GGTTGGTTCA ATTCCTTCCA TAATTCAGAA GTATAATCCA 4500
AACAATCAGG ATTATATTGA TGAATTGCCA TCATCTGATA ATCAGGAATA TGATGATAAT 4560
TCCGCTCCTT CTGGTGGTTT CTTTGTTCCG CAAAATGATA ATGTTACTCA AACTTTTAAA 4620
ATTAATAACG TTCGGGCAAA GGATTTAATA CGAGTTGTCG AATTGTTTGT AAAGTCTAAT 4680
ACTTCTAAAT CCTCAAATGT ATTATCTATT GACGGCTCTA ATCTATTAGT TGTTAGTGCA 4740
CCTAAAGATA TTTTAGATAA CCTTCCTCAA TTCCTTTCTA CTGTTGATTT GCCAACTGAC 4800
CAGATATTGA TTGAGGGTTT GATATTTGAG GTTCAGCAAG GTGATGCTTT AGATTTTTCA 4860
TTTGCTGCTG GCTCTCAGCG TGGCACTGTT GCAGGCGGTG TTAATACTGA CCGCCTCACC 4920
TCTGTTTTAT CTTCTGCTGG TGGTTCGTTC GGTATTTTTA ATGGCGATGT TTTAGGGCTA 4980
TCAGTTCGCG CATTAAAGAC TAATAGCCAT TCAAAAATAT TGTCTGTGCC ACGTATTCTT 5040
ACGCTTTCAG GTCAGAAGGG TTCTATCTCT GTTGGCCAGA ATGTCCCTTT TATTACTGGT 5100
CGTGTGACTG GTGAATCTGC CAATGTAAAT AATCCATTTC AGACGATTGA GCGTCAAAAT 5160
GTAGGTATTT CCATGAGCGT TTTTCCTGTT GCAATGGCTG GCGGTAATAT TGTTCTGGAT 5220
ATTACCAGCA AGGCCGATAG TTTGAGTTCT TCTACTCAGG CAAGTGATGT TATTACTAAT 5280
CAAAGAAGTA TTGCTACAAC GGTTAATTTG CGTGATGGAC AGACTCTTTT ACTCGGTGGC 5340

```
CTCACTGATT ATAAAAACAC TTCTCAAGAT TCTGGCGTAC CGTTCCTGTC TAAAATCCCT   5400
TTAATCGGCC TCCTGTTTAG CTCCCGCTCT GATTCCAACG AGGAAAGCAC GTTATACGTG   5460
CTCGTCAAAG CAACCATAGT ACGCGCCCTG TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT   5520
GGTTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT   5580
CTTCCCTTCC TTTCTCGCCA CGTTCGCCGG CTTTCCCCGT CAAGCTCTAA ATCGGGGGCT   5640
CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC CCCAAAAAAC TTGATTTGGG   5700
TGATGGTTCA CGTAGTGGGC CATCGCCCTG ATAGACGGTT TTTCGCCCTT TGACGTTGGA   5760
GTCCACGTTC TTTAATAGTG GACTCTTGTT CCAAACTGGA ACAACACTCA ACCCTATCTC   5820
GGGCTATTCT TTTGATTTAT AAGGGATTTT GCCGATTTCG GAACCACCAT CAAACAGGAT   5880
TTTCGCCTGC TGGGGCAAAC CAGCGTGGAC CGCTTGCTGC AACTCTCTCA GGGCCAGGCG   5940
GTGAAGGGCA ATCAGCTGTT GCCCGTCTCG CTGGTGAAAA GAAAAACCAC CCTGGCGCCC   6000
AATACGCAAA CCGCCTCTCC CCGCGCGTTG GCCGATTCAT TAATGCAGCT GGCACGACAG   6060
GTTTCCCGAC TGGAAAGCGG GCAGTGAGCG CAACGCAATT AATGTGAGTT AGCTCACTCA   6120
TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCGT ATGTTGTGTG GAATTGTGAG   6180
CGGATAACAA TTTCACACGC CAAGGAGACA GTCATAATGA AATACCTATT GCCTACGGCA   6240
GCCGCTGGAT TGTTATTACT CGCTGCCCAA CCAGCCATGG CCGAGCTCTT CCCGCCATCT   6300
GATGAGCAGT TGAAATCTGG AAGTGCCTCT GTTGTGTGCC TGCTGAATAA CTTCTATCCC   6360
AGAGAGGCCA AAGTACAGTG GAAGGTGGAT AACGCCCTCC AATCGGGTAA CTCCCAGGAG   6420
AGTGTCACAG AGCAGGACAG CAAGGACAGC ACCTACAGCC TCAGCAGCAC CCTGACGCTG   6480
AGCAAAGCAG ACTACGAGAA ACACAAAGTC TACGCCTGCG AAGTCACCCA TCAGGGCCTG   6540
AGCTCGCCCG TCACAAAGAG CTTCAACAGG GGAGAGTGTT CTAGAACGCG TCACTTGGCA   6600
CTGGCCGTCG TTTTACAACG TCGTGACTGG GAAAACCCTG GCGTTACCCA AGCTTAATCG   6660
CCTTGCAGAA TTCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA CCGATCGCCC   6720
TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCTTT GCCTGGTTTC CGGCACCAGA   6780
AGCGGTGCCG GAAAGCTGGC TGGAGTGCGA TCTTCCTGAG GCCGATACGG TCGTCGTCCC   6840
CTCAAACTGG CAGATGCACG GTTACGATGC GCCCATCTAC ACCAACGTAA CCTATCCCAT   6900
TACGGTCAAT CCGCCGTTTG TTCCCACGGA GAATCCGACG GGTTGTTACT CGCTCACATT   6960
TAATGTTGAT GAAAGCTGGC TACAGGAAGG CCAGACGCGA ATTATTTTTG ATGGCGTTCC   7020
TATTGGTTAA AAAATGAGCT GATTTAACAA AAATTTAACG CGAATTTTAA CAAAATATTA   7080
ACGTTTACAA TTTAAATATT TGCTTATACA ATCTTCCTGT TTTTGGGGCT TTTCTGATTA   7140
TCAACCGGGG TACATATGAT TGACATGCTA GTTTTACGAT TACCGTTCAT CGATTCTCTT   7200
GTTTGCTCCA GACTCTCAGG CAATGACCTG ATAGCCTTTG TAGATCTCTC AAAAATAGCT   7260
ACCCTCTCCG GCATTAATTT ATCAGCTAGA ACGGTTGAAT ATCATATTGA TGGTGATTTG   7320
ACTGTCTCCG GCCTTTCTCA CCCTTTTGAA TCTTTACCTA CACATTACTC AGGCATTGCA   7380
```

TTTAAAATAT ATGAGGGTTC TAAAAATTTT TATCCTTGCG TTGAAATAAA GGCTTCTCCC 7440

GCAAAAGTAT TACAGGGTCA TAATGTTTTT GGTACAACCG ATTTAGCTTT ATGCTCTGAG 7500

GCTTTATTGC TTAATTTTGC TAATTCTTTG CCTTGCCTGT ATGATTTATT GGATGTT 7557

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8118 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: both
(D) TOPOLOGY: circular

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AATGCTACTA CTATTAGTAG AATTGATGCC ACCTTTTCAG CTCGCGCCCC AAATGAAAAT 60

ATAGCTAAAC AGGTTATTGA CCATTTGCCA AATGTATCTA ATGGTCAAAC TAAATCTACT 120

CGTTCGCAGA ATTGGGAATC AACTGTTACA TCGAATGAAA CTTCCAGACA CCGTACTTTA 180

GTTGCATATT TAAAACATGT TGAGCTACAG CACCAGATTC AGCAATTAAG CTCTAAGCCA 240

TCTGCAAAAA TGACCTCTTA TCAAAAGGAG CAATTAAAGG TACTCTCTAA TCCTGACCTG 300

TTGGAGTTTG CTTCCGGTCT GGTTCGCTTT GAAGCTCGAA TTAAAACGCG ATATTTGAAG 360

TCTTTCGGGC TTCCTCTTAA TCTTTTTGAT GCAATCCGCT TTGCTTCTGA CTATAATAGT 420

CAGGGTAAAG ACCTGATTTT TGATTTATGG TCATTCTCGT TTTCTGAACT GTTTAAAGCA 480

TTTGAGGGGG ATTCAATGAA TATTTATGAC GATTCCGCAG TATTGGACGC TATCCAGTCT 540

AAACATTTTA CTATTACCCC CTCTGGCAAA ACTTCTTTTG CAAAAGCCTC TCGCTATTTT 600

GGTTTTTATC GTCGTCTGGT AAACGAGGGT TATGATAGTG TTGCTCTTAC TATGCCTCGT 660

AATTCCTTTT GGCGTTATGT ATCTGCATTA GTTGAATGTG GTATTCCTAA ATCTCAACTG 720

ATGAATCTTT CTACCTGTAA TAATGTTGTT CCGTTAGTTC GTTTTATTAA CGTAGATTTT 780

TCTTCCCAAC GTCCTGACTG GTATAATGAG CCAGTTCTTA AAATCGCATA AGGTAATTCA 840

CAATGATTAA AGTTGAAATT AAACCATCTC AAGCCCAATT TACTACTCGT TCTGGTGTTT 900

CTCGTCAGGG CAAGCCTTAT TCACTGAATG AGCAGCTTTG TTACGTTGAT TTGGGTAATG 960

AATATCCGGT TCTTGTCAAG ATTACTCTTG ATGAAGGTCA GCCAGCCTAT GCGCCTGGTC 1020

TGTACACCGT TCATCTGTCC TCTTTCAAAG TTGGTCAGTT CGGTTCCCTT ATGATTGACC 1080

GTCTGCGCCT CGTTCCGGCT AAGTAACATG GAGCAGGTCG CGGATTTCGA CACAATTTAT 1140

CAGGCGATGA TACAAATCTC CGTTGTACTT TGTTTCGCGC TTGGTATAAT CGCTGGGGGT 1200

CAAAGATGAG TGTTTTAGTG TATTCTTTCG CCTCTTTCGT TTTAGGTTGG TGCCTTCGTA 1260

GTGGCATTAC GTATTTTACC CGTTTAATGG AAACTTCCTC ATGAAAAAGT CTTTAGTCGT 1320

CAAAGCCTCT GTAGCCGTTG CTACCCTCGT TCCGATGCTG TCTTTCGCTG CTGAGGGTGA 1380

CGATCCCGCA AAAGCGGCCT TTAACTCCCT GCAAGCCTCA GCGACCGAAT ATATCGGTTA 1440

TGCGTGGGCG ATGGTTGTTG TCATTGTCGG CGCAACTATC GGTATCAAGC TGTTTAAGAA 1500

ATTCACCTCG AAAGCAAGCT GATAAACCGA TACAATTAAA GGCTCCTTTT GGAGCCTTTT 1560

TTTTTGGAGA TTTTCAACGT GAAAAAATTA TTATTCGCAA TTCCTTTAGT TGTTCCTTTC 1620

TATTCTCACT CCGCTGAAAC TGTTGAAAGT TGTTTAGCAA AACCCCATAC AGAAAATTCA 1680

TTTACTAACG TCTGGAAAGA CGACAAAACT TTAGATCGTT ACGCTAACTA TGAGGGTTGT 1740

CTGTGGAATG CTACAGGCGT TGTAGTTTGT ACTGGTGACG AAACTCAGTG TTACGGTACA 1800

TGGGTTCCTA TTGGGCTTGC TATCCCTGAA AATGAGGGTG GTGGCTCTGA GGGTGGCGGT 1860

TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT ACTAAACCTC CTGAGTACGG TGATACACCT 1920

ATTCCGGGCT ATACTTATAT CAACCCTCTC GACGGCACTT ATCCGCCTGG TACTGAGCAA 1980

AACCCCGCTA ATCCTAATCC TTCTCTTGAG GAGTCTCAGC CTCTTAATAC TTTCATGTTT 2040

CAGAATAATA GGTTCCGAAA TAGGCAGGGG GCATTAACTG TTTATACGGG CACTGTTACT 2100

CAAGGCACTG ACCCCGTTAA AACTTATTAC CAGTACACTC CTGTATCATC AAAAGCCATG 2160

TATGACGCTT ACTGGAACGG TAAATTCAGA CACTGCGCTT TCCATTCTGG CTTTAATGAA 2220

GATCCATTCG TTTGTGAATA TCAAGGCCAA TCGTCTGACC TGCCTCAACC TCCTGTCAAT 2280

GCTGGCGGCG GCTCTGGTGG TGGTTCTGGT GGCGGCTCTG AGGGTGGTGG CTCTGAGGGT 2340

GGCGGTTCTG AGGGTGGCGG CTCTGAGGGA GGCGGTTCCG GTGGTGGCTC TGGTTCCGGT 2400

GATTTTGATT ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT 2460

GAAAACGCGC TACAGTCTGA CGCTAAAGGC AAACTTGATT CTGTCGCTAC TGATTACGGT 2520

GCTGCTATCG ATGGTTTCAT TGGTGACGTT TCCGGCCTTG CTAATGGTAA TGGTGCTACT 2580

GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG GTGACGGTGA TAATTCACCT 2640

TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA ATGTCGCCCT 2700

TTTGTCTTTA GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA 2760

TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG 2820

TTTGCTAACA TACTGCGTAA TAAGGAGTCT TAATCATGCC AGTTCTTTTG GGTATTCCGT 2880

TATTATTGCG TTTCCTCGGT TTCCTTCTGG TAACTTTGTT CGGCTATCTG CTTACTTTTC 2940

TTAAAAAGGG CTTCGGTAAG ATAGCTATTG CTATTTCATT GTTTCTTGCT CTTATTATTG 3000

GGCTTAACTC AATTCTTGTG GGTTATCTCT CTGATATTAG CGCTCAATTA CCCTCTGACT 3060

TTGTTCAGGG TGTTCAGTTA ATTCTCCCGT CTAATGCGCT TCCCTGTTTT TATGTTATTC 3120

TCTCTGTAAA GGCTGCTATT TTCATTTTTG ACGTTAAACA AAAAATCGTT TCTTATTTGG 3180

ATTGGGATAA ATAATATGGC TGTTTATTTT GTAACTGGCA AATTAGGCTC TGGAAAGACG 3240

CTCGTTAGCG TTGGTAAGAT TCAGGATAAA ATTGTAGCTG GGTGCAAAAT AGCAACTAAT 3300

CTTGATTTAA GGCTTCAAAA CCTCCCGCAA GTCGGGAGGT TCGCTAAAAC GCCTCGCGTT 3360

CTTAGAATAC CGGATAAGCC TTCTATATCT GATTTGCTTG CTATTGGGCG CGGTAATGAT 3420

TCCTACGATG AAAATAAAAA CGGCTTGCTT GTTCTCGATG AGTGCGGTAC TTGGTTTAAT 3480

ACCCGTTCTT GGAATGATAA GGAAAGACAG CCGATTATTG ATTGGTTTCT ACATGCTCGT 3540

AAATTAGGAT GGGATATTAT TTTTCTTGTT CAGGACTTAT CTATTGTTGA TAAACAGGCG 3600
CGTTCTGCAT TAGCTGAACA TGTTGTTTAT TGTCGTCGTC TGGACAGAAT TACTTTACCT 3660
TTTGTCGGTA CTTTATATTC TCTTATTACT GGCTCGAAAA TGCCTCTGCC TAAATTACAT 3720
GTTGGCGTTG TTAAATATGG CGATTCTCAA TTAAGCCCTA CTGTTGAGCG TTGGCTTTAT 3780
ACTGGTAAGA ATTTGTATAA CGCATATGAT ACTAAACAGG CTTTTTCTAG TAATTATGAT 3840
TCCGGTGTTT ATTCTTATTT AACGCCTTAT TTATCACACG GTCGGTATTT CAAACCATTA 3900
AATTTAGGTC AGAAGATGAA GCTTACTAAA ATATATTTGA AAAAGTTTTC ACGCGTTCTT 3960
TGTCTTGCGA TTGGATTTGC ATCAGCATTT ACATATAGTT ATATAACCCA ACCTAAGCCG 4020
GAGGTTAAAA AGGTAGTCTC TCAGACCTAT GATTTTGATA AATTCACTAT TGACTCTTCT 4080
CAGCGTCTTA ATCTAAGCTA TCGCTATGTT TTCAAGGATT CTAAGGGAAA ATTAATTAAT 4140
AGCGACGATT TACAGAAGCA AGGTTATTCA CTCACATATA TTGATTTATG TACTGTTTCC 4200
ATTAAAAAAG GTAATTCAAA TGAAATTGTT AAATGTAATT AATTTTGTTT TCTTGATGTT 4260
TGTTTCATCA TCTTCTTTTG CTCAGGTAAT TGAAATGAAT AATTCGCCTC TGCGCGATTT 4320
TGTAACTTGG TATTCAAAGC AATCAGGCGA ATCCGTTATT GTTTCTCCCG ATGTAAAAGG 4380
TACTGTTACT GTATATTCAT CTGACGTTAA ACCTGAAAAT CTACGCAATT TCTTTATTTC 4440
TGTTTTACGT GCTAATAATT TTGATATGGT TGGTTCAATT CCTTCCATAA TTCAGAAGTA 4500
TAATCCAAAC AATCAGGATT ATATTGATGA ATTGCCATCA TCTGATAATC AGGAATATGA 4560
TGATAATTCC GCTCCTTCTG GTGGTTTCTT TGTTCCGCAA AATGATAATG TTACTCAAAC 4620
TTTTAAAATT AATAACGTTC GGGCAAAGGA TTTAATACGA GTTGTCGAAT TGTTTGTAAA 4680
GTCTAATACT TCTAAATCCT CAAATGTATT ATCTATTGAC GGCTCTAATC TATTAGTTGT 4740
TAGTGCACCT AAAGATATTT TAGATAACCT TCCTCAATTC CTTTCTACTG TTGATTTGCC 4800
AACTGACCAG ATATTGATTG AGGGTTTGAT ATTTGAGGTT CAGCAAGGTG ATGCTTTAGA 4860
TTTTTCATTT GCTGCTGGCT CTCAGCGTGG CACTGTTGCA GGCGGTGTTA ATACTGACCG 4920
CCTCACCTCT GTTTTATCTT CTGCTGGTGG TTCGTTCGGT ATTTTTAATG GCGATGTTTT 4980
AGGGCTATCA GTTCGCGCAT TAAAGACTAA TAGCCATTCA AAAATATTGT CTGTGCCACG 5040
TATTCTTACG CTTTCAGGTC AGAAGGGTTC TATCTCTGTT GGCCAGAATG TCCCTTTTAT 5100
TACTGGTCGT GTGACTGGTG AATCTGCCAA TGTAAATAAT CCATTTCAGA CGATTGAGCG 5160
TCAAAATGTA GGTATTTCCA TGAGCGTTTT TCCTGTTGCA ATGGCTGGCG GTAATATTGT 5220
TCTGGATATT ACCAGCAAGG CCGATAGTTT GAGTTCTTCT ACTCAGGCAA GTGATGTTAT 5280
TACTAATCAA AGAAGTATTG CTACAACGGT TAATTTGCGT GATGGACAGA CTCTTTTACT 5340
CGGTGGCCTC ACTGATTATA AAAACACTTC TCAAGATTCT GGCGTACCGT TCCTGTCTAA 5400
AATCCCTTTA ATCGGCCTCC TGTTTAGCTC CCGCTCTGAT TCCAACGAGG AAAGCACGTT 5460
ATACGTGCTC GTCAAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTA AGCGCGGCGG 5520
GTGTGGTGGT TACGCGCAGC GTGACCGCTA CACTTGCCAG CGCCCTAGCG CCCGCTCCTT 5580

TCGCTTTCTT CCGTTCGTTT GTCGCCACGT TCGCCGGCTT TCCCCGTCAA GCTCTAAATC 5640
GGGGGCTCCC TTTAGGGTTC CGATTTAGTG CTTTACGGCA CCTCGACCCC AAAAAACTTG 5700
ATTTGGGTGA TGGTTCACGT AGTGGGCCAT CGCCCTGATA GACGGTTTTT CGCCCTTTGA 5760
CGTTGGAGTC CACGTTCTTT AATAGTGGAC TCTTGTTCCA AACTGGAACA ACACTCAACC 5820
CTATCTCGGG CTATTCTTTT GATTTATAAG GGATTTTGCC GATTTCGGAA CCACCATCAA 5880
ACAGGATTTT CGCCTGCTGG GGCAAACCAG CGTGGACCGC TTGCTGCAAC TCTCTCAGGG 5940
CCAGGCGGTG AAGGGCAATC AGCTGTTGCC CGTCTCGCTG GTGAAAAGAA AAACCACCCT 6000
GGCGCCCAAT ACGCAAACCG CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC 6060
ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC 6120
TCACTCATTA GGCACCCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA 6180
TTGTGAGCGG ATAACAATTT CACACGCCAA GGAGACAGTC ATAATGAAAT ACCTATTGCC 6240
TACGGCAGCC GCTGGATTGT TATTACTCGC TGCCCAACCA GCCATGGCCG AGCTCTTCCC 6300
GCCATCTGAT GAGCAGTTGA AATCTGGAAC TGCCTCTGTT GTGTGCCTGC TGAATAACTT 6360
CTATCCCAGA GAGGCCAAAG TACAGTGGAA GGTGGATAAC GCCCTCCAAT CGGGTAACTC 6420
CCAGGAGAGT GTCACAGAGC AGGACAGCAA GGACAGCACC TACAGCCTCA GCAGCACCCT 6480
GACGCTGAGC AAAGCAGACT ACGAGAAACA CAAAGTCTAC GCCTGCGAAG TCACCCATCA 6540
GGGCCTGAGC TCGCCCGTCA CAAAGAGCTT CAACAGGGGA GAGTGTTCTA GAACGCGTCA 6600
CTTGGCACTG GCCGTCGTTT TACAACGTCG TGACTGGGAA AACCCTGGCG TTACCCAAGC 6660
TTTGTACATG GAGAAAATAA AGTGAAACAA AGCACTATTG CACTGGCACT CTTACCGTTA 6720
CTGTTTACCC CTGTGGCAAA AGCCGCCTCC ACCAAGGGCC CATCGGTCTT CCCCCTGGCA 6780
CCCTCCTCCA AGAGCACCTC TGGGGGCACA GCGGCCCTGG GCTGCCTGGT CAAGACTAAT 6840
TCCCCGAACC GGTGACGGTG TCGTGGAACT CAGGCGCCCT GACCAGCGGC GTGCACACCT 6900
TCCCGGCTGT CCTACAGTCC TCAGGACTCT ACTCCCTCAG CAGCGTGGTG ACCGTGCCCT 6960
CCAGCAGCTT GGGCACCCAG ACCTACATCT GCAACGTGAA TCACAAGCCC AGCAACACCA 7020
AGGTGGACAA GAAAGCAGAG CCCAAATCTT GTACTAGTGG ATCCTACCCG TACGACGTTC 7080
CGGACTACGC TTCTTAGGCT GAAGGCGATG ACCCTGCTAA GGCTGCATTC AATAGTTTAC 7140
AGGCAAGTGC TACTGAGTAC ATTGGCTACG CTTGGGCTAT GGTAGTAGTT ATAGTTGGTG 7200
CTACCATAGG GATTAAATTA TTCAAAAAGT TTACGAGCAA GGCTTCTTAA GCAATAGCGA 7260
AGAGGCCCGC ACCGATCGCC CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGCGCTT 7320
TGCCTGGTTT CCGGCACCAG AAGCGGTGCC GGAAAGCTGG CTGGAGTGCG ATCTTCCTGA 7380
GGCCGATACG GTCGTCGTCC CCTCAAACTG GCAGATGCAC GGTTACGATG CGCCCATCTA 7440
CACCAACGTA ACCTATCCCA TTACGGTCAA TCCGCCGTTT GTTCCCACGG AGAATCCGAC 7500
GGGTTGTTAC TCGCTCACAT TTAATGTTGA TGAAAGCTGG CTACAGGAAG GCCAGACGCG 7560
AATTATTTTT GATGGCGTTC CTATTGGTTA AAAAATGAGC TGATTTAACA AAAATTTAAC 7620

GCGAATTTTA ACAAAATATT AACGTTTACA ATTTAAATAT TTGCTTATAC AATCTTCCTG 7680

TTTTTGGGGC TTTTCTGATT ATCAACCGGG GTACATATGA TTGACATGCT AGTTTTACGA 7740

TTACCGTTCA TCGATTCTCT TGTTTGCTCC AGACTCTCAG GCAATGACCT GATAGCCTTT 7800

GTAGATCTCT CAAAAATAGC TACCCTCTCC GGCATTAATT TATCAGCTAG AACGGTTGAA 7860

TATCATATTG ATGGTGATTT GACTGTCTCC GGCCTTTCTC ACCCTTTTGA ATCTTTACCT 7920

ACACATTACT CAGGCATTGC ATTTAAAATA TATGAGGGTT CTAAAAATTT TTATCCTTGC 7980

GTTGAAATAA AGGCTTCTCC CGCAAAAGTA TTACAGGGTC ATAATGTTTT TGGTACAACC 8040

GATTTAGCTT TATGCTCTGA GGCTTTATTG CTTAATTTTG CTAATTCTTT GCCTTGCCTG 8100

TATGATTTAT TGGACGTT 8118

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: misc_difference
(B) LOCATION: replace(5, "")
(D) OTHER INFORMATION: /note= "S REPRESENTS EQUAL MIXTURE OF G AND C"

(ix) FEATURE:
(A) NAME/KEY: misc_difference
(B) LOCATION: replace(6, "")
(D) OTHER INFORMATION: /note= "M REPRESENTS EQUAL MIXTURE OF A AND C"

(ix) FEATURE:
(A) NAME/KEY: misc_difference
(B) LOCATION: replace(8, "")
(D) OTHER INFORMATION: /note= "R REPRESENTS EQUAL MIXTURE OF A AND G"

(ix) FEATURE:
(A) NAME/KEY: misc_difference
(B) LOCATION: replace(11, "")
(D) OTHER INFORMATION: /note= "K REPRESENTS EQUAL MIXTURE OF G AND T"

(ix) FEATURE:
(A) NAME/KEY: misc_difference
(B) LOCATION: replace(20, "")
(D) OTHER INFORMATION: /note= "W REPRESENTS EQUAL MIXTURE OF A AND T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGGTSMARCT KCTCGAGTCW GG 22

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AGGTCCAGCT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

AGGTCCAGCT GCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGGTCCAGCT TCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGGTCCAGCT TCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: lin ar

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AGGTCCAACT GCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGGTCCAACT GCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AGGTCCAACT TCTCGAGTCT GG 22

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AGGTCCAACT TCTCGAGTCA GG 22

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:
(A) NAME/KEY: misc_difference
(B) LOCATION: replace(5..6, "")
(D) OTHER INFORMATION: /note= "N-INOSINE"

(ix) FEATURE:
(A) NAME/KEY: misc_difference
(B) LOCATION: replace(8, "")
(D) OTHER INFORMATION: /note= "N-INOSINE"

(ix) FEATURE:
(A) NAME/KEY: misc_difference
(B) LOCATION: replace(11, "")
(D) OTHER INFORMATION: /note= "N-INOSINE"

(ix) FEATURE:
(A) NAME/KEY: misc_difference
(B) LOCATION: replace(20, "")
(D) OTHER INFORMATION: /note= "W REPRESENTS EQUAL MIXTURE OF A AND T"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AGGTNNANCT NCTCGAGTCW GG 22

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTATTAACTA GTAACGGTAA CAGTGGTGCC TTGCCCCA 38

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGGCTTACTA GTACAATCCC TGGGCACAAT 30

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCAGTTCCGA GGTCGTTGTG ACTCAGGAAT CT 32

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 bas pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: lin ar

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCAGTTCCGA GCTCGTGTTG ACGCAGCCGC CC 32

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CCAGTTCCGA GCTCGTGCTC ACCCAGTCTC CA 32

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CCAGTTCCGA GCTCCAGATG ACCCAGTCTC CA 32

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CCAGATGTGA GCTCGTGATG ACCCAGACTC CA 32

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CCAGATGTGA GCTCGTCATG ACCCAGTCTC CA 32

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singl
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CCAGTTCCGA GCTCGTGATG ACACAGTCTC CA 32

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCAGCATTCT AGAGTTTCAG CTCCAGCTTG CC 32

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCGCCGTCTA GAATTAACAC TCATTCCTGT TGAA 34

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GATCCTAGGC TGAAGGCGAT GACCCTGCTA AGGCTGC 37

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

ATTCAATAGT TTACAGGCAA GTGCTACTGA GTACA 35

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TTGGCTACGC TTGGGCTATG GTAGTAGTTA TAGTT 35

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GGTGCTACCA TAGGGATTAA ATTATTCAAA AAGTT 35

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TACGAGCAAG GCTTCTTA 18

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

AGCTTAAGAA GCCTTGCTCG TAAACTTTTT GAATAATTT 39

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AATCCCTATG GTAGCACCAA CTATAACTAC TACCAT 36

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

AGCCCAAGCG TAGCCAATGT ACTCAGTAGC ACTTG 35

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

CCTGTAAACT ATTGAATGCA GCCTTAGCAG GGTC 34

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

ATCGCCTTCA GCCTAG 16

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

CATTTTTGCA GATGGCTTAG A 21

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

TAGCATTAAC GTCCAATA 18

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

ATATATTTTA GTAAGCTTCA TCTTCT 26

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GACAAAGAAC GCGTGAAAAC TTT 23

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GCGGGCCTCT TCGCTATTGC TTAAGAAGCC TTGCT 35

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

AAACGACGGC CAGTGCCAAG TGACGCGTGT GAAATTGTTA TCC 43

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GGCGAAAGGG AATTCTGCAA GGCGATTAAG CTTGGGTAAC GCC 43

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GGCGTTACCC AAGCTTTGTA CATGGAGAAA ATAAAG 36

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

TGAAACAAAG CACTATTGCA CTGGCACTCT TACCGTTACC GT 42

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

TACTGTTTAC CCCTGTGACA AAAGCCGCCC AGGTCCAGCT GC 42

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 44 base pairs
(B) TYPE: nucl ic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

TCGAGTCAGG CCTATTGTGC CCAGGGATTG TACTAGTGGA TCCG 44

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

TGGCGAAAGG GAATTCGGAT CCACTAGTAC AATCCCTG 38

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

GGCACAATAG GCCTGACTCG AGCAGCTGGA CCACGGCGGC TT 42

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

TTGTCACAGG GGTAAACAGT AACGGTAACG GTAAGTGTGC CA 42

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GTGCAATAGT GCTTTGTTTC ACTTTATTTT CTCCATGTAC AA 42

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singl
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

TAACGGTAAG AGTGCCAGTG C 21

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

CACCTTCATG AATTCGGCAA GGAGACAGTC AT 32

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

AATTCGCCAA GGAGACAGTC AT 22

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

AATGAAATAC CTATTGCCTA CGGCAGCCGC TGGATTGTT 39

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

ATTACTCGCT GCCCAACCAG CCATGGCCGA GCTCGTGAT 39

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

GACCCAGACT CCAGATATCC AACAGGAATG AGTGTTAAT 39

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

TCTAGAACGC GTC 13

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

TTCAGGTTGA AGCTTACGCG TTCTAGAATT AACACTCATT CCTGT 45

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

TGGATATCTG GAGTCTGGGT CATCACGAGC TCGGCCATG 39

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singl
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

GCTGGTTGGG CAGCGAGTAA TAACAATCCA GCGGCTGCC 39

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

GTAGGCAATA GGTATTTCAT TATGACTGTC CTTGGCG 37

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

TGACTGTCTC CTTGGCGTGT GAAATTGTTA 30

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

TAACACTCAT TCCGGATGGA ATTCTGGAGT CTGGGT 36

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

GCCAGTGCCA AGTGACGCGT TCTA 24

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

ATATATTTTA GTAAGCTTCA TCTTCT 26

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

GACAAAGAAC GCGTGAAAAC TTT 23

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 76 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CTGAACCTGT CTGGGACCAC AGTTGATGCT ATAGGATCAG ATCTAGAATT CATTTAGAGA 60

CTGGCCTGGC TTCTGC 76

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

TCGACCGTTG GTAGGAATAA TGCAATTAAT GGAGTAGCTC TAAATTCAGA ATTCATCTAC 60

ACCCAGTGCA TCCAGTAGCT 80

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

GGTAAACAGT AACGGTAAGA GTGCCAG 27

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 54 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

CGCCTTCAGC CTAAGAAGCG TAGTCCGGAA CGTCGTACGG GTAGGATCCA CTAG 54

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CACCGGTTCG GGGAATTAGT CTTGACCAGG CAGCCCAGGG C 41

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

ATTCCACACA TTATACGAGC CGGAAGCATA AAGTGTCAAG CCTGGGGTGC C 51

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CTGCTCATCA GATGGCGGGA AGAGCTCGGC CATGGCTGGT TG 42

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singl
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

GAACAGAGTG ACCGAGGGGG CGAGCTCGGC CATGGCTGGT TG

42